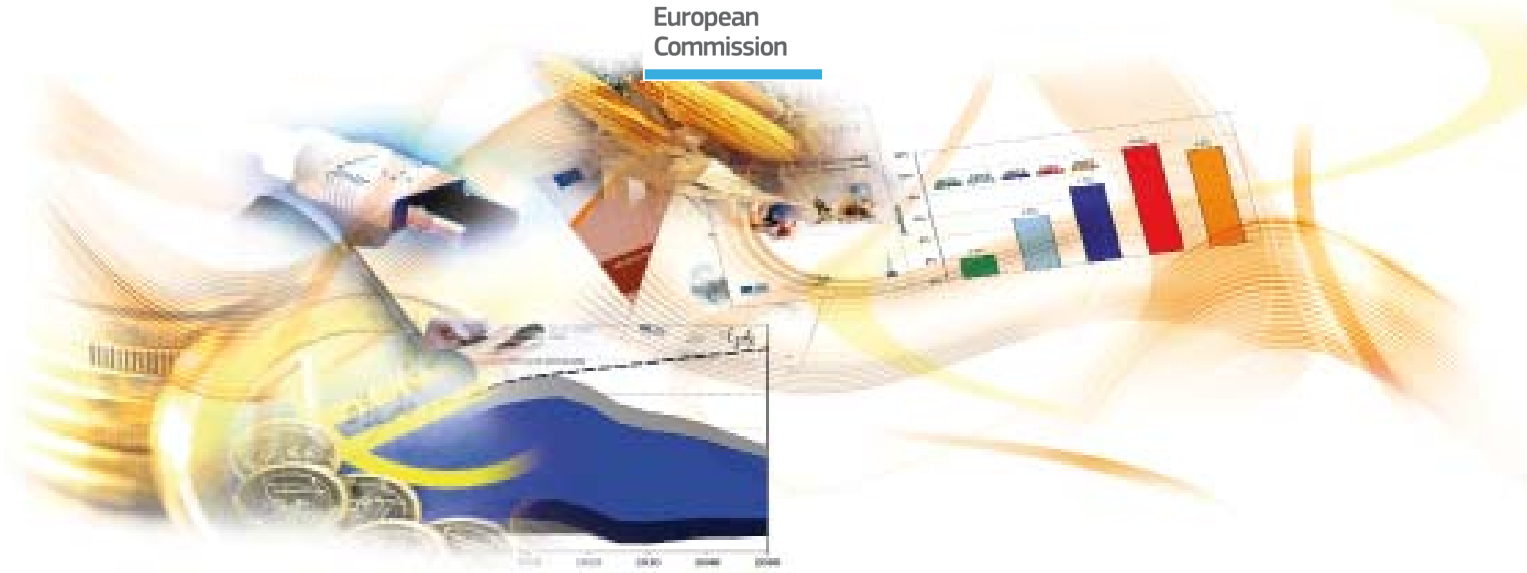
European Commission



J R C T E C H N I C A L R E P O R T S

Institute for Prospective Technological Studies
Digital Economy Working Paper 2013/06

Has the Digital Divide Been Reversed? Evidence from Five EU Countries

Smaranda Pantea

Bertin Martens

2013

Report EUR 25959 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Address: Edificio Expo. C/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-secretariat@ec.europa.eu
Tel.: +34 954488318
Fax: +34 954488300

http://ipts.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

This publication is a Working Paper by the Joint Research Centre of the European Commission. It results from the Digital Economy Research Programme at the JRC Institute for Prospective Technological Studies, which carries out economic research on information society and EU Digital Agenda policy issues, with a focus on growth, jobs and innovation in the Single Market. The Digital Economy Research Programme is co-financed by the Directorate General Communications Networks, Content and Technology.

JRC80782

EUR 25959 EN

ISBN 978-92-79-29822-6 (pdf)

ISSN 1831-9424 (online)

doi:10.2791/12705

Luxembourg: Publications Office of the European Union, 2013

Printed in Spain

Abstract[1]

This paper examines whether there is a digital divide in the use of the internet in general and for specific purposes (leisure, improving human capital and obtaining goods and services). It uses a unique dataset which covers the entire clickstream of almost 20,000 internet users in the five largest EU economies during 2011. Our main finding is that, for those who have access to the Internet, the income-based digital divide in internet use has been reversed. Low-income internet users spend more time on the internet than high-income users. In addition, we find that employment status does not change the effect of income on internet use and we discuss several possible explanations for this result. There is some evidence of an education-based digital divide in the use of human capital and goods & services websites. Tertiary education has a negative effect on time spent on leisure websites and a positive effect on time spent on human capital and goods & services websites. Using quantile regressions, we find that the negative effect of income and the positive effect of education for human capital and goods & services websites hold for the entire conditional distribution of these online activities. Moreover, these effects are stronger for more intensive internet users.

JEL codes: *L86, D12, D13*

Keywords: *Internet Use, Time allocation, Leisure.*

[1] Valuable comments and suggestions from Russel Cooper, Pierre Montaigner and Marc Bogdanowicz are gratefully acknowledged. We remain responsible for all the errors and omissions.

1. Introduction

There is considerable policy concern about the digital divide between people with access to the internet and those without. This divide is often evident in differences in socio- economic characteristics, especially income and education. The digital divide in access to internet has been extensively documented. However, less is known about the digital divide in internet use for people who have access to the internet, especially in Europe, and where it exists it is based mainly on survey data (Demoussis and Giannakopolous, 2006; Orviska and Hudson, 2009; Montagnier and Wirthmann, 2011). This paper aims to contribute to this area by studying the digital divide in internet use in the five largest countries in the EU. More precisely, it studies how, for those who have access to the Internet, income and education affect internet use in general and its use for specific purposes (leisure, improving human capital and obtaining goods and services). It uses a unique dataset that covers the entire clickstream of almost 20,000 internet users in the five largest EU economies during the year 2011.

This paper builds on Goldfarb and Prince (2008) who study the role of income and education levels in internet use patterns in the US. We extend this study in several ways. First, we study the determinants of time spent online in the five largest EU economies using data on internet users' online behaviour (their entire clickstream), which is more objective and precise than survey data. Second, we study the determinants of time spent on three specific types of websites: (a) human capital improvement, such as career, education and health-related sites, (b) obtaining goods and services and (c) leisure. Third, we study whether the effects of income and education on time spent online differ by employment status (employed or not employed) and by intensity of internet use.

Our main finding is that, for those who have access to the Internet, the digital divide is reversed: low-income users spend more time online than high-income users. This relationship is particularly strong for time spent on leisure online. Internet users in the lowest income range also spent most time on websites related to human capital and obtaining goods and services, however for users above the lowest income group there is no relationship between time spent online on these types of websites and household income. We find that education has a positive effect on time spent on websites related to human capital and to obtaining services and goods, which points towards an education-related digital divide in the use of these websites. Somewhat surprisingly in the view of opportunity cost hypothesis, we find that employment status (working or not working) does not affect the effect of income on internet use. Using quantile regressions, we find that the negative effect of income on time spent on internet and the positive effect for

time spent on human capital and goods & services websites hold for the entire conditional distribution of these online activities. Moreover these effects are stronger for more intensive internet users.

The paper is organised as it follows. Section 2 reviews the related literature. Section 3 describes the data used and presents some preliminary evidence on the relationship between time spent online and income and other demographic characteristics. Section 4 describes the empirical methodology. Section 5 discusses the results of the estimation and Section 6 offers conclusions.

2. Related Literature

Internet use has been studied from several angles and a review of all the literature on this topic is beyond the purpose of this paper. The paper is related to three strands of literature: studies on the welfare effects and value of internet use, studies related to use of internet for a specific purpose and sociological studies on the effect of internet use on other activities.

Our study is mostly related to studies that examine internet use measured as time spent online and its welfare effects (Goolsbee and Klenow, 2006; Goldfarb and Prince, 2008; Brynjolfsson and Oh, 2012). In these studies, in line with Becker (1965), consumer utility depends on his consumption, which requires income obtained partly through labour, and on his leisure. Individuals choose the time spent on online leisure, offline leisure and work in order to maximise their utility subject to budget and time constraints (the sum of the amount of time spent daily on each of these activities cannot be higher than 24 hours). Internet pricing consists of fixed cost of adoption and zero use fees. Then, conditional on internet adoption, the marginal cost of internet use is given only by the opportunity cost of time, which is determined by the income that the internet user could earn on the labour market. Therefore, the opportunity cost of spending time online is higher for high income earners than for low income ones. If both low and high income users benefit equally from internet, than given the higher opportunity costs for high income users, they will spend less time on internet.

Goolsbee and Klenow (2006), Goldfarb and Prince (2008) and Brynjolfsson and Oh (2012) test empirically this hypothesis on a sample of US internet users. All these studies find that, conditional on internet access, income has a negative effect on time spent online. Most of these studies suggest that higher opportunity costs of spending time online of the high income internet users explain this negative relationship. However, Goldfarb and Prince (2008) notices that there are several alternative explanations which are consistent with this negative relationship: (1) the opportunity cost hypothesis explained above (2) internet is more useful for low-income people,

because they have different preferences or because they do not have better offline alternatives (3) low income earners have more leisure time, which leads them to spend more time online even if they have the same opportunity costs as high income people (4) cost of adoption of internet is an important cost barrier for low income individuals, but not for higher income and therefore only low income earners which place a higher value on internet will adopt internet, but most of high income people will adopt internet (including those who do not place a high value on internet). It is important to mention these explanations do not exclude each other. The authors find that the selection and the amount of leisure time are not driving their results, but they find some evidence of differences in usefulness of internet for different income groups. They conclude that the main explanation of the negative relationship between time spent online and income is the higher opportunity cost of time for high income users, but that there is some evidence that internet is more useful for low income people. We explore more in detail these two explanations.[2]

There are several studies that examine the use of the internet for specific purposes such as: e-commerce, job search, entertainment etc. (Demoussis and Giannakopolous, 2006; Goldfarb and Prince, 2008; Orviska and Hudson, 2009; Montagnier and Wirthmann, 2011; Pérez-Hernández and Sánchez-Mangas, 2011). Due to differences in data sources and in the dependent and explanatory variables they are not directly comparable. However, they show some common patterns. The most important and most relevant for our study is that income, education, and other demographic characteristics have different effects on participation in different online activities. Goldfarb and Prince (2008) find that for US internet users, income and university/college education is negatively associated with using internet for activities related to leisure (chat, online games) and e-health, but positively associated with using it for activities related to buying (research purchases and ecommerce). Pérez-Hernández and Sánchez-Mangas (2011) found similar effects of education on online shopping for Spain. Demoussis and Giannakopoulos (2006) find that education and household income are important determinants of the frequency of internet use in Europe and Montagnier and Wirthmann (2011) find that they are important determinants of using internet daily. Although we draw on these studies, we differ from them in that our study does not examine the determinants of using internet for a specific purpose or with a certain frequency, but of the time spent on different online activities.

[2] We cannot explore the selection hypothesis (4) because our sample consists entirely of active internet users. We will study differences across countries. We also drop hypothesis (3) because we have no information on total leisure time available to internet users. However, we control for it through proxy variables such as users' occupation and life stage.

Our paper is also related to sociology studies that examine which activities are displaced by time spent on internet. The findings of this literature are mixed and overall suggest that spending time on internet is not consistently associated with notable changes in media use or social activities and other daily activities (for a review see Martin and Robinson (2010)). A more nuanced view is proposed by Nie and Hillygus (2002). They suggest that only intensive internet use (more than 60 minutes per day) has a large effect on other activities, especially on leisure and to a lesser extent on work, childcare, housework and sleeping, while light internet use has a small and often insignificant effect. These findings suggest that income should have a higher effect on time spent online for intensive internet users, than for light users. We will test this hypothesis in the empirical part of the paper.

In conclusion, there is a large and very heterogeneous literature related to the topic of internet use, including a few studies on the relationship between income and/or education and internet use. However, most of these empirical studies are based on US survey data and most of them do not take into account several aspects of this relationship documented in other strands of literature (different types of online activities, intensity of use, and differences across different demographic groups). In this paper, we study different aspects of the effect of income and education on internet use using objective clickstream data from five large EU countries.

3. Data Description

The data used in this paper have been collected by Nielsen NetRatings through voluntary online consumer panels. The dataset contains information on all web pages clicked on from their home computers by 25,000 internet users in the five largest EU economies (France, Germany, Italy, Spain and United Kingdom) during the entire year 2011. According to Nielsen, the sample of internet users is representative of the *online* population in these countries in terms of gender and age[3]. For each click it contains information on the URL, the time and date the website is accessed and time spent on the website. The data on the online activity is collected through a piece of software that internet users in the online panel voluntarily install on their PC. The data collection procedure uses information in the computer about which webpage is in focus (the page to which the keyboard and mouse activity is directed to). This helps correct for errors in measurement of the time spent on websites due to minimising tabs, tabbed browsing and periods of inactivity.

[3] Nielsen provides incentives to participate and to remain in the panel in the form of vouchers and points which can be redeemed from their reward website or used in online games and sweepstakes(prize drawing), which might bias our sample towards people who are more likely to value these activities. In the empirical part we will discuss potential implications for our estimated effects and as a robustness check we will repeat the estimations excluding time spent on online games and gambling websites to make sure that our results are not driven by time spent on these websites.

Nielsen also corrects for some of the main potential problems that could lead to errors in measurement of the traffic (error pages, aborted pages, pages not request by the PC, websites which include frames, etc.). For most websites the dataset contains their brand names, which are classified into subcategories and categories based on the content of the websites using a methodology developed by Nielsen. For each user[4] the dataset contains information on basic social and economic characteristics, gathered through a questionnaire when the user installs the Nielsen software.

The sample that we use in the empirical analysis excludes records with missing information on the website category and on the demographic characteristics of the internet user, and records of unlikely young and old[5] internet users and outliers[6] and records of self-employed internet users.[7] Table 1 shows how excluding these observations affects the sample used for the empirical analysis. The remaining samples are still large covering close to 4,000 users in each country, which represent close to 80% of the initial user sample and more than 70% of the initial clickstream sample.

We examine how much time users spend online and on what type of websites (leisure, human capital related and related to buying/obtaining goods or services). Table 2 present the classification of websites into these groups of online activities, which is based on how each activity contributes to consumer utility, in line with Becker (1965) and Gronau (1977): leisure (contributes directly to the utility), human capital websites, such as career, health and education (contributes to the utility through future income which can be spent of future consumption) and goods & services websites (contributes to the utility as an input in the production of the final goods/services consumed).

---

4 There are households in which more users are registered with Nielsen. In these households the meter prompts the internet user to log in; however the match between user profile and his online activity is likely to be imperfect. To ensure that our results are not affected by this problem, we will estimate our model also on the sample of one person households.

5 The dataset includes internet users between 2 and 99 years old. It is likely that very young and very old consumers did not answer the questionnaire themselves and that they did not use the internet themselves. To ensure comparability we will focus on internet users aged between 16 and 74 in line with Eurostat for Information Society Indicators and previous empirical studies (Orviska and Hudson, 2009, Perez-Hernandez and Sanchez-Mangas, 2011).

6 We exclude internet users who spend an implausible large or small amount of time online (internet users in the highest and the lowest 1% of average weekly time spent online). The main results are not affected by this exclusion.

7 In the case of self employed we do not know what part of their time online is related to their work and which part is leisure. They represent 9% of our sample, but the results are robust to including these observations in the sample.

Table 4 presents the average time spent on each of the three groups of online activities. The average person spends 5 hours per week online: close to 3 hours and half on leisure websites, more than one hour on goods & services websites and around 8 minutes per week on websites related to work, education and health. The most popular leisure websites are, in order, social networks, online games, videos/movies and adult websites, the most popular goods & services websites are general portals, search and e-commerce websites and the most popular human capital websites are career websites.

We examine the social and economic characteristics of the internet users in the sample. The definition of all variables used and their summary statistics for the internet users in the sample aged 16-74 - excluding outliers and those with missing information - are given in Table 3 and Table 4. The summary statistics of the demographic characteristics of the internet users show that the sample used in the empirical analysis includes a large variety of internet users in terms of education, occupation, income and other demographic characteristics.

Figure 1 presents the distribution of time spent online (minutes per week) in the pooled 5-country sample on all types of websites combined and on the three specific types that we identify: leisure, human capital and obtaining goods and services. It shows the percentage of internet users in the sample on the vertical axis that spent a specific number of minutes (a multiple of 60 minutes for all activities except human capital for which we use a multiple of 10 minutes) online per week on the horizontal axis. The figure reveals large heterogeneity in the intensity of using internet. Many internet users use spent little time online and at there is long tail that spends many hours online.

In Figure 2, we present some patterns that show how time spent online is linked with income. It shows that total time spent online and time spent on all online activities considered decreases with income. This relationship is strongest for all time spent online and for time spent on leisure websites, for the other two types of websites it is weaker. These patterns show clearly that the income-based digital divide in internet use is reversed: low-income users spend more time online than high-income users, for all three types of websites. These patterns are consistent with the hypothesis that high income users have a higher opportunity cost of time and therefore spend less time on these online activities.

Figure 3 presents how average time spent online varies with educational attainment. Internet users with tertiary education spent less time online than users with lower educational attainment. This pattern might indicate higher opportunity cost of time for internet users with tertiary

education. There is a clear positive relationship between human capital websites and time spent online, which indicates that there might be a digital divide in the ability to use these websites according to education levels. Finally, there is no relationship between time spent online on goods & services websites and education.

In summary, our descriptive analysis shows that there is large heterogeneity regarding the time spent online and that there is a negative relationship between income and time spent online and mixed relationship between education and time spent online.

4. Methodology

Following Goldfarb and Prince (2008), we assume time spent online is a function of total leisure time, total income, price of internet and other individual characteristics. We include controls for occupational and demographic characteristics related to life stage (being married/cohabitating and having children) to control for leisure time. Household income is our proxy for total money available. We include country and region dummies to control for possible differences in the fixed cost of an internet connection. We also include several demographic characteristics which previous studies have shown to have an effect on time spent online. We estimate the following regression:

$$y_i = \alpha + \beta_I' Income_i + \beta_E' Education_i + \beta_x' x_i + \beta_o' o_i + \beta_c' c_i + \beta_{cr}' r_{ci} + u_i \qquad (1)$$

y_i is the average time spent online per week by internet user i, measured in minutes per week. Since we do not have a continuous income variable but only income groups, $Income_i$ is measured as a set of dummies for household income in a given interval. $Eduction_i$ are dummy variables that control for the educational level of the internet user. o_i are dummy variables for different occupations, c_i and r_{ci} are dummy variables that control for country and regions within each country where the internet user resides. x_i are other social and economic characteristics of the internet user. Informed by previous empirical studies on the topic, we also include the following control variables: gender, age, being single, children in the household and household size. The exact definitions for all variables are in Table 3 and descriptive statistics are in Table 4.

The main variables of interest are *Income* and *Education*. The opportunity cost of time hypothesis,[8] predicts that the opportunity cost of spending time online is higher for high income earners. Consequently, they will spend less time online overall and possibly also on different types of

[8] We do not include a financial cost in the opportunity cost in this model because we assume fixed monthly internet use fees and consequently a zero marginal financial cost of internet use.

internet activities. Finding negative and significant coefficients on the income dummies ($Income_i$) will be interpreted as confirming this hypothesis. However, it is important to mention that other effects such as preferences, network effects, and other effects may affect these coefficients.

In this model, education is a proxy for the ability of using internet for different purposes. Previous studies show there is a strong relationship between education and digital skills[9]. If an online activity requires certain skills/abilities we would expect a positive coefficient of education. We would expect this to be true for more sophisticated internet uses, especially human capital and goods & services websites, which includes e-commerce and the use of different online services such as online banking or government websites[10], but not necessarily for most of leisure activities. Education may affect time spent online also through opportunity cost of spending time online and preferences. To the extent that highly educated internet users earn higher wages they also face higher opportunity costs of time. Controlling for income should account for this effect. In addition, high and low educated individuals may differ in their preferences regarding different online activities. However, in the empirical part we will look at aggregated groups of activities and also at a large number of detailed online activities, which would allow us to examine whether education is more relevant to more complex online activities or not.

We test whether the effects of income and education on time spent online vary across the distribution of time spent online on different online activities. Testing this hypothesis is relevant given the positively skewed distribution of time spent online (Figure 1). Given that OLS estimations are sensitive to outliers, it is important to confirm that our results are not driven by a few very intensive internet users. Nie and Hillygus (2002) find that intensive internet use crowds out other daily activities, while light use of internet does not. This suggests the effect of income may vary across the conditional distribution of time spent online. Ability (education) might also have a different effect on light or intensive internet users. It could be that light users only use basic digital content which does not require much ability, while more intensive users use more diverse and sophisticated features of the websites and digital content, which requires higher ability or digital skills.

---

9 In the countries included in the study in 2011, after learning by doing, formal education is the second most important way in which individuals acquire digital skills (Eurostat Information Society Indicators, 2013). In addition, there is a large literature that documents the relationship between education and adoption and use of ICT (Autor, Katz and Krueger, 1998).

10 In Netherlands, CBS (2008) found that education level is an important determinant of shopping online using e-government services, especially downloading and returning completed forms and documents.

Equation (1) is estimated using OLS. In our sample all individuals have access to internet at home and were active users of internet during the period studied. Therefore, we cannot examine the determinants of adoption of the internet or control for it. Moreover, the descriptive statistics presented in the previous section show that all users spent positive amounts of time on leisure and goods & services websites and 98.6% do so on human capital websites. Given that our dependent variable is not censored, or in the case of human capital websites it is be very little affected by censoring, we conclude that OLS is the appropriate estimation method.[11] In addition, we will carry out several robustness checks which will be discussed in the results section.

To test whether the effects of income and education are robust to the presence of outliers and whether they vary across the distribution of internet use we will use quantile regressions. This method provides a more complete characterisation of the conditional distribution of time spent online by allowing the effect of income and education and other explanatory variable that to vary and it is more appropriate given the positively skewed distribution of our dependent variables (see Figure 1). Quantile regression estimation is not sensitive to outliers and it is more robust and efficient than OLS when the error is non normal (Buchinsky, 1998).

Most of the regressions are estimated on the pooled sample of all five countries excluding outliers, observations with missing information and self employed internet users described in Section 3. However, we will report country specific results for the baseline model. We found the same relationships between income and education and time spent online in all countries (although the magnitudes of the effects vary slightly) and therefore we decided to focus on the pooled sample.

5. Estimation Results

Table 5 reports the estimation of equation (1). The results in the first column of Table 5 confirm that all income coefficients are negative and statistically significant. The household income group 0-18,000 Euros is taken as the reference group. *Ceteris paribus*, internet users in the second lowest household income group (18,000 - 27,000 Euros) spend on average 50 minutes per week less online than users in the lowest income group (less than 18,000); users in the highest income group (above 72.000) spend 2.5 hours per week less online. The differences between the coefficients of income intervals are statistically significant (see bottom part of Table 5). The results suggest that time spent online decreases almost monotonically with the household

[11] We have estimated the equation (1) using tobit and the results are very similar to the ones obtained using OLS.

income. First, these results are consistent with the opportunity cost hypothesis and with previous studies (Goolsbee and Klenow, 2006; Goldfarb and Prince, 2008; Brynjolfsson and Oh, 2012). Second, they suggest that there is a reverse in the income-based digital divide in internet use: low-income users use internet more than high-income users.

Columns 3, 4 and 5 in Table 5 show that this negative relationship between income and time spent holds for each type of website considered. For leisure, the differences between the coefficients for the income intervals are statistically significant, suggesting a monotonically increasing negative income effect. This is consistent with the opportunity cost hypothesis, and with previous empirical studies on leisure online (Goolsbee and Klenow, 2006; Goldfarb and Prince, 2008; Brynjolfsson and Oh, 2012). It is also similar to results on TV watching (Frey *et al.*, 2007). The differences between income coefficients are not statistically significant for time spent on human capital and goods & services websites (reported in the bottom part of Table 5). This suggests that for internet users above the lowest income interval there is no relationship between time spent online on human capital and goods & services websites.

Secondary education has a positive effect on time spent on all types of websites, although this effect is insignificant for time spent on leisure websites. Tertiary education has an insignificant effect on overall time spent online, a negative effect on time spent on leisure websites, and a significant and positive effect on time spent on human capital and goods & services websites. These results are consistent with Goldfarb and Prince's (2008) findings that education has a negative effect on time spent on leisure, but a positive effect on e-commerce and research prior to purchases. The results for education on time spent on human capital and goods & services websites suggests there might be a divide in terms of ability to use these types of websites.

The other variables included have the expected signs. Large household size is associated with less times spent online, while being single is associated with more time spent on all online activities. The coefficient for the variable controlling for the presence of children is negative, but mostly insignificant due to the high correlation with household size variables. Female internet users spend less time on leisure websites, but more time on human capital and goods& services websites. Age has a mixed impact. On all categories of websites except human capital websites, internet users that have a technical or professional job spend least time online and internet users who are unemployed or homemakers spend most time online. Internet users who are unemployed, students or work in education spend most time on human capital websites.

These estimations are based on aggregated groups of activities. Aggregation of different types of websites that correspond to online activities may result in smoothing the effect of income and education. Therefore, we repeat the estimations at a more disaggregated level of categories of websites. These results are reported in Table 6. Leisure websites comprise entertainment and lifestyle, news, social networks and internet services websites, which include communication services, such as email, instant messaging, long distance call and downloading websites. The income coefficients remain negative and statistically significant, similar to those of the more aggregate leisure category, except for time spent on news websites, for which they are insignificant. Education coefficients remain negative for entertainment and social networks, but turn positive and significant for news and internet services. Human capital categories include career and education, corporate and health websites. Coefficients for both income and educations are similar to those of the aggregate category human capital, but income coefficients are not always significant. Goods & services categories include ecommerce, general portal and search, travel, online banking, government & non-profit websites. Income coefficients vary significantly for these categories. They are negative and significant for ecommerce and general portal & search, but insignificant for government and non-profit websites and positive and significant for travel and online banking. The positive coefficients for these two categories suggests that these last categories are used more by higher income users, either due to preferences or pure income effects (the effect of an increase in income when the prices, including opportunity costs, remain constant). The coefficients of education are always positive and significant for each of the goods & services categories. Overall, these results are in line with those obtained from more aggregated categories, but there are a few categories for which they differ. The positive and significant effect of education on all categories in human capital and goods & services categories and more sophisticated entertainment categories such as news suggest that this effect is related to ability.

We carry out several robustness checks. To address possible problems with the measurement of duration of time spent online, we re-estimate equation (1) replacing time spent with the number of clicks per week as the dependant variable. The results of these estimations are reported in Table 7 and confirm the results for time spent online. As indicated in footnote 4, the match between user profile and online activity may not be perfect in households with several individuals. Therefore, as a robustness check we estimate the baseline model on the sample of one person households, which are not affected by this problem. These results are shown in Table 8 and they are qualitatively similar to the baseline results, although the magnitude of the

coefficients is different. However, it is important to mention that one person households differ significantly from other households in terms of sociological characteristics. We also estimate equation (1) excluding time spent on online games and gambling/sweepstakes to check that our results are not driven by time spent on these websites. The results in Table 9 confirm the baseline results for income and education.

In Table 10 to Table 13 we show country specific results to confirm that our results are not driven by one country or a group of countries. The results show that in all countries there is a negative and significant relationship between time spent online and income and a positive relationship between education and time spent on human capital websites. Overall these country results show that there is no income-based digital divide in internet use in any of the five countries, but that in all countries there is an education-based divide in the use of websites related to career, education and health and websites related to obtaining goods and services.

We also test whether the effects of income and education on time spent online varies by employment status of internet users. This test is useful for two purposes. First, it tests whether the negative relationship between income and internet is driven by access to internet at work for higher income internet users. Second, it provides an additional test for the opportunity cost of time hypothesis. The results (reported in Table 14) show that income has a negative effect on overall time use and on time spent on specific websites, both for working and non-working users. Moreover, the differences in coefficients between working and not-working internet users are not statistically significant. These results indicate that the negative relationship between time spent online and income are not driven by the fact that high income internet users a more likely to have access to internet at work than low income ones. However, the absence of any statistically significant difference between the income coefficients for working and not-working internet users is more difficult to interpret and may cast doubt over the opportunity cost of working time hypothesis. The opportunity cost of time of internet users who do not earn a work-related income is not given by income earned in the labour market, but by other possible uses of their time. For instance, the opportunity costs of time for students may be given by time spent on education (and thus earning future income), for homemakers by time spent on childcare or home work and for unemployed by time spent on searching for a job. However, these opportunity costs may be correlated with household income. These results could also be interpreted as lending supports to the hypothesis that low income internet users benefit more from internet than high income internet users who may have better alternatives or different preferences as suggested by Goldfarb and Prince (2008). Our model does not allow us to distinguish between

these possible explanations. A further research step would be to specify a model which allows doing this.

We check whether the wide heterogeneity and long-tail distribution of the observations affects the findings regarding income and other explanatory variables. OLS estimates are sensitive to the presence of outliers and therefore we want to test whether our results for income and education are not driven by a few very intensive users (users that spend large amounts of time online). In addition, the online activity of the very intensive internet users is itself of interest because they account for a large part of online activity. For instance, the top 10% internet users in the distribution of the online activities studied account for 36% of time spent on goods & services websites, for 40% of time spent on leisure online and for more than 50% of total time spent on human capital websites. For this purpose we use quantile regressions. The estimation results for the 10th, 25th, 50th, 75th and 90th quantiles, for the four categories of websites, are reported in Table 15- Table 18.

The results show that the coefficients of income variables are significant across the entire conditional distribution of time spent online for all four activities considered. This confirms that our OLS results are not driven by outliers. The effect of income is higher at the top than at the bottom of the distribution of time spent online. We also tested and confirmed that the differences in the income coefficient for different quantiles are statistically significant. These results suggest that income has a greater effect for more intensive users. These results are consistent with the hypothesis that heavy use of internet crowds out other activities and light use does not (Nie and Hillygus, 2002), but given the limits of our model it is difficult to test this specific hypothesis.

Education has different effects for different online activities. For overall time and leisure, education has a positive effect on the lower quantiles of these distributions but an insignificant and even negative effect on the higher quantiles of these distributions. For time spent on human capital and goods & services websites, education has a positive and significant effect of the distributions of these online activities, confirming our OLS results. Its effect is higher for higher quantiles of these distributions. These results could be interpreted as evidence of the need for skills to use these types of websites intensively; leisure sites apparently require less skills or only basic skills relevant at very low levels of use.

6. Conclusions

This paper aims to contribute to the debate regarding digital divide in access and use of internet between individuals with different socio-economic characteristics, especially income and education. While there is a large amount of literature on the digital divide in access, less is known about the digital divide in use. The evidence that exists is based on survey data for US.

We build on Goldfarb and Prince (2008), who study the role of income and education levels on internet use patterns in the US, which we extend to study the determinants of three specific online activities: leisure, human capital improvement and obtaining goods and services. In addition, we study whether the relationship between income and education and time spent online differs between employed and not employed internet users (who may differ in their access to internet and their opportunity costs of time), and between light and heavy internet users.

Our main finding is that, for those who have access to internet, the income based digital divide in internet use has been reversed: low income internet users spend more time online overall and on websites related to leisure. Internet users in the lowest income group also spend more time on human capital and goods & services websites. However, we find evidence that there may be an education-related digital divide in the use of human capital and goods & services websites, which are generally regarded as valuable online activities and policymakers seek to encourage.

These results are robust to several robustness checks. Employment status does not change the effects of income and education. These effects hold for the entire conditional distribution of the online activities considered. The effect of income on time spent on all the online activities considered is stronger for more intensive users. The effect of education on time spent on human capital and goods & services websites is also stronger for more intensive users of these websites.

Overall the results suggest that for those with access to internet there is a reversal in the income-based digital divide in internet use. They suggest that currently the main digital divide in internet use is driven by education and it concerns not the internet in general, but specific uses such as the use of human capital and obtaining goods and services (including e-government, online banking services and online shopping). The results highlight the importance of education for enabling internet users to participate in online activities generally regarded as valuable.

References

Autor, D. H., Katz L. F. and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, vol. 113(4), 1169-1213.

Brynjolfsson, E. and J. H. Oh (2012):"The Attention Economy: Measuring the Value of Free Goods on the Internet", mimeo.

Becker, G. S. (1965): "A Theory of the Allocation of Time," *Economic Journal,* vol. 75 (299), 493-517.

Buchinsky, M (1998): "Recent Advances in Quantile Regression Models: A Practical Guide for Empirical Research," *Journal of Human Resources*, vol. 33(1), 88 -126.

CBS (2008): The Digital Economy 2008, Central Bureau of Statistics –Statistics Netherlands.

Demoussis, M. and N. Giannakopoulos (2006): "Facets of the Digital Divide in Europe: Determination and Extent of Internet Use," *Economics of Innovation and New Technology*, vol. 15(3), 235-246.

Eurostat (2010): Eurostat Information Society Indicators Database.

Frey, B. S., Benesch C., and A. Stutzer (2007): "Does Watching TV Makes Us Happy?", *Journal of Economic Psychology*, vol. 28(3), 283-313.

Goldfarb, A. and J.Prince (2008): "Internet Adoption and Usage Patterns Are Different: Implications for the Digital Divide," *Information Economics and Policy*, vol. 20(1), 2-15.

Goolsbee, G and Klenow, P. J. (2006): "Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet," *American Economic Review*, vol. 96(2), 108-113.

Gronau, R. (1977): "Leisure, Home Production, and Work-The Theory of the Allocation of Time Revisited," *Journal of Political Economy*, vol. 85(6), 1099-1123.

Montagnier, P and A. Wirthmann (2011): "Digital Divide: from Computer Access to Online Activities- A Micro Data Analysis," OECD Digital Economy Working Papers no. 189, OECD.

Nie, N. D. and D. S. Hillygus (2002): "The Impact of Internet Use on Sociability: Time-Diary Findings," *IT & Society*, vol. 1(1), 1-20.

OECD (2012): Internet Economy Outlook 2012.

Orviska, M. and J. Hudson (2009): "Dividing or Uniting Europe? Internet Usage in the EU," *Information Economics and Policy*, vol. 21(4), 279-290.

Pérez-Hernández, J. and R. Sánchez-Mangas (2011): " To Have or Not to Have Internet at Home: Implications for Online Shopping," *Information Economics and Policy*, vol. 23(3), 213-226.

Robinson, S. M and J. P. Marin (2010): "IT Use and Declining Social Capital? More Cold Water from the General Social Survey (GSS) and the American Time Use Survey (ATUS)," *Social Science Computer Review*, vol. 28(1), 45-63.

Annex

Tables

Table 1: Summary of available information about clicks.

Country	Total Clicks		Categorised clicks		Categorised clicks with complete demographic data		Categorised clicks of consumers aged 16-74 years, with complete demographic data and not considered outliers	
1	2	3	4	5	6	7	8	9
	users	clicks	share users	share clicks	share users	share clicks	share users	share clicks
FR	5000	147.601.904	100%	89%	96%	85%	83%	74%
DE	5000	246.568.640	100%	87%	93%	81%	86%	73%
IT	5000	211.011.296	100%	91%	90%	81%	82%	73%
ES	5000	222.590.768	100%	90%	91%	83%	83%	75%
UK	5000	199.006.384	100%	90%	84%	78%	76%	70%

Source: Nielsen Click stream

Table 2: Online Activity Definition

Activity	Nielsen Category
All Websites	All categories of websites
Leisure Websites	Entertainment, Family and Lifestyle (except subcategory Health, Nutrition and Fitness), News & Information, subcategories Member Communities and Targeted Member Communities from Portals & Communities category and Internet Services.
Human Capital Websites	Education& Careers, Corporate,[12] subcategory Health, Nutrition and Safety from Family and Lifestyle).
Goods & Services Websites	Home & Fashion, Ecommerce, Travel, Government & Nonprofit, Finance, Search Engines, General Portals & Search (subcategories General Portals and Search from Search Engines, Portals & Communities category), Special Occasions, Automotive, Computers & Electronics.

[12] We assume that people searched this category mainly for finding information about job vacancies. However, classifying it as a residual category or as goods & services webiste does not change the results.

Table 3: Variable definitions

Variable	Definition
All Websites	Average time spent online per week (in minutes)
Leisure Websites	Average time spent online per week on leisure websites (in minutes).
Human Capital Websites	Average time spent online per week on human capital websites (in minutes)
Goods & Services Websites	Average time spent online per week on goods & services websites (in minutes)
All Websites Clicks	Average number of clicks per week.
Leisure Websites Clicks	Average number of clicks per week on leisure websites.
Human Capital Websites Clicks	Average number of clicks per week on human capital websites
Goods & Services Websites Clicks	Average number of clicks per week on goods & services websites related
Female	Dummy variable indicating whether the r is female.
Single	Dummy variable indicating whether the internet user is single (not married or cohabitating).
Age	Age of the internet users (in years)
Children	Dummy variable indicating the presence of children.
Household size 1-2	Dummy variable indicated that there are 1 or 2 people in the household of the internet user.
Household size 3-4	Dummy variable indicated that there are 3 or 4 people in the household of the internet user.
Household size>5	Dummy variable indicated that there are more than 5 people in the household of the internet user
Income ≤18000	Dummy variable indicating whether the internet user' household income is in this interval.
Income 18-27000	Dummy variable indicating whether the internet user' household income is in this interval.
Income 27-36001	Dummy variable indicating whether the internet user' household income is in this interval.
Income 36-54000	Dummy variable indicating whether the internet user' household income is in this interval.
Income 54-72000	Dummy variable indicating whether the internet user' household income is in this interval.
Income >72000	Dummy variable indicating whether the internet user' household income is in this interval.
Below Secondary Education	Dummy variable indicating whether the highest educational attainment of the internet user is below secondary education.
Secondary Education	Dummy variable indicating whether the highest educational attainment of the internet user is secondary education.
Tertiary Education	Dummy variable indicating whether the highest educational attainment of the internet user is tertiary education.
Employed	Dummy variable indicating whether the internet user is employed.
Clerical/Administrative	Dummy variable indicating whether the internet user has a clerical/administrative job.
Craftsman/Craftswoman	Dummy variable indicating whether the internet user has a clerical/administrative job.
Education	Dummy variable indicating whether the internet user has an education job.
Executive/Managerial	Dummy variable indicating whether the internet user has an executive/managerial job.
Military	Dummy variable indicating whether the internet user has a military job.
Operator/Labourer	Dummy variable indicating whether the internet user has an operator/labourer job.
Other	Dummy variable indicating whether the internet user has other job.
Professional	Dummy variable indicating whether the internet user has a professional job.
Sales	Dummy variable indicating whether the internet user has sales job.
Service	Dummy variable indicating whether the internet user has service job.
Technical	Dummy variable indicating whether the internet user has technical job.
Unemployed	Dummy variable indicating whether the internet user is unemployed.
Student	Dummy variable indicating whether the internet user is a student.
Retired	Dummy variable indicating whether the internet user is retired.
Homemaker	Dummy variable indicating whether the internet user is a homemaker or carer.

Source: Calculations based on Nielsen Click stream

Table 4: Summary statistics

Variable	Obs.	Mean	Std.	Min	Max
All Websites	18680	306.02	320.30	1.78	1973.38
Leisure Websites	18680	214.05	270.18	0.02	1838.78
Human Capital Websites	18680	8.79	18.53	0.00	556.86
Goods & Services Websites	18680	83.18	93.62	0.04	1479.06
All Websites Clicks	18680	697.15	889.64	1.37	11161.60
Leisure Websites Clicks	18680	482.83	758.66	0.08	11028.79
Human Capital Websites Clicks	18680	20.01	51.60	0.00	2630.19
Goods & Services Websites Clicks	18680	194.31	242.47	0.17	3562.83
Female	18680	0.51	0.50	0	1
Single	18680	0.25	0.43	0	1
Age	18680	41.68	13.57	16	74
Children	18680	0.31	0.46	0	1
Household size 1-2	18680	0.52	0.50	0	1
Household size 3-4	18680	0.41	0.49	0	1
Household size >5	18680	0.08	0.27	0	1
Income ≤18000	18680	0.21	0.41	0	1
Income 18-27000	18680	0.23	0.42	0	1
Income 27-36001	18680	0.17	0.38	0	1
Income 36-54000	18680	0.22	0.42	0	1
Income 54-72000	18680	0.10	0.30	0	1
Income >72000	18680	0.06	0.23	0	1
Below Secondary Education	18680	0.26	0.44	0	1
Secondary Education	18680	0.26	0.44	0	1
Tertiary Education	18680	0.48	0.50	0	1
Employed	18680	0.66	0.47	0	1
Clerical/Administrative	18680	0.17	0.37	0	1
Craftsman/Craftswoman	18680	0.01	0.11	0	1
Education	18680	0.04	0.20	0	1
Executive/Managerial	18680	0.09	0.29	0	1
Military	18680	0.01	0.09	0	1
Operator/Labourer	18680	0.07	0.25	0	1
Other	18680	0.06	0.24	0	1
Professional	18680	0.04	0.20	0	1
Sales	18680	0.04	0.19	0	1
Service	18680	0.07	0.25	0	1
Technical	18680	0.06	0.24	0	1
Unemployed	18680	0.09	0.29	0	1
Student	18680	0.09	0.29	0	1
Retired	18680	0.10	0.29	0	1
Homemaker	18680	0.07	0.25	0	1

Source: Calculations based on Nielsen Click stream

Table 5: Baseline model

	All Websites	Leisure Websites	Human Capital Websites	Goods & Services Websites
Income 18-27000	-50.48	-43.04	-1.62	-5.82
	[7.80]***	[6.75]***	[0.42]***	[2.16]***
Income 27-36001	-65.76	-59.91	-1.64	-4.20
	[8.21]***	[7.04]***	[0.46]***	[2.31]*
Income 36-54000	-99.47	-88.55	-2.14	-8.78
	[7.87]***	[6.68]***	[0.46]***	[2.31]***
Income 54-72000	-122.48	-107.49	-2.52	-12.47
	[9.11]***	[7.55]***	[0.60]***	[2.85]***
Income >72000	-148.03	-123.43	-3.80	-20.80
	[10.41]***	[8.57]***	[0.63]***	[3.14]***
Secondary educ.	17.71	6.34	2.33	9.03
	[7.23]**	[6.18]	[0.37]***	[2.01]***
Tertiary educ.	-1.85	-16.66	3.39	11.43
	[6.75]	[5.75]***	[0.35]***	[1.94]***
Female	-1.84	-8.48	2.77	3.87
	[4.85]	[4.10]**	[0.30]***	[1.45]***
Age	-1.18	-1.59	0.04	0.37
	[0.24]***	[0.20]***	[0.01]***	[0.07]***
Single	76.89	59.48	1.86	15.54
	[6.81]***	[5.84]***	[0.42]***	[2.01]***
Children	-7.75	-7.54	-0.95	0.74
	[6.41]	[5.35]	[0.38]**	[1.86]
Household size 3-4	-30.31	-22.37	-0.43	-7.51
	[6.83]***	[5.68]***	[0.42]	[2.07]***
Household size >5	-32.53	-22.10	-0.83	-9.60
	[10.64]***	[8.82]**	[0.60]	[3.18]***
Constant	398.29	334.40	5.77	58.11
	[21.19]***	[18.05]***	[1.15]***	[5.65]***
N	18680	18680	18680	18680
R^2	0.10	0.10	0.03	0.05
Mean of Y	306.02	214.05	8.79	83.18
F tests of differences in income coefficients (p values)				
$\beta_{Inc.\ 18\text{-}27000} = \beta_{Inc.\ 27\text{-}36000}$	0.03	0.00	0.94	0.47
$\beta_{Inc.\ 27\text{-}36000} = \beta_{Inc.\ 36\text{-}54000}$	0.00	0.00	0.21	0.02
$\beta_{Inc.\ 36\text{-}54000} = \beta_{Inc.\ 54\text{-}72000}$	0.00	0.00	0.46	0.11
$\beta_{Inc.\ 54\text{-}72000} = \beta_{Inc.\ >72000}$	0.00	0.02	0.04	0.01

Notes: Dependent variable is average time spent per week on all or on a specified category of websites, measured in minutes. All equations include occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 6: Detailed website categories

	Leisure				Human Capital			Goods & Services				
	Entertainment & Lifestyle	News	Social Networks	Internet Services	Careers & Education	Corporate	Health	Ecommerce	Search & General Portals	Travel	Online Banking	E-Gov. & Nonprofit
Income 18-27000	-16.42	0.49	-22.41	-4.71	-4.20	0.59	-2.11	-0.06	-0.05	-0.97	-0.51	-0.14
	[3.49]***	[0.81]	[4.09]***	[1.25]***	[1.50]***	[0.25]**	[0.72]***	[0.41]	[0.20]	[0.28]***	[0.25]**	[0.10]
Income 27-36001	-20.46	-0.90	-32.37	-6.18	-3.09	1.09	-2.66	0.39	0.08	-0.93	-0.58	-0.14
	[3.68]***	[0.71]	[4.18]***	[1.34]***	[1.59]*	[0.27]***	[0.75]***	[0.44]	[0.23]	[0.32]***	[0.28]**	[0.09]
Income 36-54000	-32.91	0.07	-46.10	-9.61	-6.82	1.19	-4.32	1.10	0.08	-1.30	-0.77	-0.08
	[3.45]***	[0.73]	[3.97]***	[1.27]***	[1.57]***	[0.25]***	[0.73]***	[0.59]*	[0.23]	[0.33]***	[0.26]***	[0.11]
Income 54-72000	-42.68	-0.59	-50.69	-13.54	-9.46	1.12	-5.41	1.31	-0.03	-2.00	-0.48	-0.04
	[3.72]***	[0.90]	[4.51]***	[1.51]***	[1.87]***	[0.34]***	[0.99]***	[0.66]**	[0.24]	[0.40]***	[0.40]	[0.12]
Income >72000	-52.80	-0.13	-52.24	-18.26	-13.92	1.41	-8.22	0.35	-0.41	-2.45	-1.23	-0.12
	[3.95]***	[1.20]	[5.30]***	[1.66]***	[1.98]***	[0.42]***	[0.99]***	[0.73]	[0.28]	[0.46]***	[0.31]***	[0.18]
Secondary educ.	-1.69	3.36	0.76	3.92	3.54	1.30	2.21	1.28	0.70	1.18	0.90	0.25
	[3.24]	[0.66]***	[3.67]	[1.08]***	[1.40]**	[0.23]***	[0.66]***	[0.39]***	[0.14]***	[0.22]***	[0.25]***	[0.10]***
Tertiary educ.	-12.47	4.89	-16.40	7.32	3.28	1.65	2.97	2.39	1.12	2.63	0.49	0.27
	[3.12]***	[0.64]***	[3.28]***	[1.09]***	[1.35]**	[0.20]***	[0.63]***	[0.45]***	[0.14]***	[0.23]***	[0.20]**	[0.09]***
Constant	121.09	9.21	183.91	20.19	22.65	3.95	23.82	4.32	3.38	4.58	1.34	-0.15
	[8.91]***	[1.96]***	[10.96]***	[3.28]***	[3.62]***	[0.65]***	[1.87]***	[1.51]***	[0.73]***	[0.87]***	[0.63]**	[0.18]
N	18680	18680	18680	18680	18680	18680	18680	18680	18680	18680	18680	18680
R^2	0.07	0.03	0.10	0.05	0.05	0.01	0.02	0.06	0.04	0.04	0.03	0.05

Notes: Dependent variable is average time spent per week on a specified category of websites, measured in minutes. All equations include occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 7: Robustness checks: Clicks

	All Websites	Leisure Websites	Human Capital Websites	Goods & Services Websites
Income 18-27000	-112.24	-95.87	-4.34	-12.04
	[21.82]***	[19.02]***	[1.24]***	[5.56]**
Income 27-36001	-139.19	-128.17	-4.41	-6.62
	[23.68]***	[20.61]***	[1.25]***	[6.09]
Income 36-54000	-222.31	-199.64	-5.32	-17.34
	[22.51]***	[19.52]***	[1.25]***	[5.96]***
Income 54-72000	-269.12	-234.43	-6.24	-28.44
	[25.91]***	[21.88]***	[1.60]***	[7.24]***
Income >72000	-349.42	-291.44	-9.88	-48.09
	[27.93]***	[23.66]***	[1.68]***	[7.92]***
Secondary educ.	53.70	24.57	5.85	23.28
	[20.45]***	[17.60]	[1.22]***	[5.44]***
Tertiary educ.	-17.41	-49.35	7.54	24.41
	[19.10]	[16.45]***	[0.94]***	[5.19]***
Constant	1124.78	911.08	19.86	193.84
	[60.11]***	[51.24]***	[3.19]***	[15.69]***
Mean of Y	697.15	482.83	20.01	194.31
N	18680	18680	18680	18680
R^2	0.08	0.08	0.02	0.03

Notes: Dependent variable is the average number of clicks per week on all or on a specific category of websites. All equations include controls for other demographic characteristics and occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 8: Robustness checks: one person households

	All Websites	Leisure Websites	Human Capital Websites	Goods & Services Websites
Income 18001-27000	-24.42	-18.26	-1.90	-4.26
	[14.70]*	[13.14]	[0.74]**	[3.90]
Income 27001-36001	-57.84	-51.12	-2.34	-4.38
	[16.14]***	[14.23]***	[0.86]***	[4.66]
Income 36001-54000	-75.65	-71.01	-0.75	-3.89
	[16.02]***	[14.00]***	[1.03]	[4.52]
Income 54001-72000	-97.17	-94.91	-1.37	-0.89
	[20.48]***	[17.51]***	[1.68]	[6.67]
Income >72000	-123.72	-106.89	-4.13	-12.70
	[24.63]***	[20.97]***	[1.27]***	[7.03]*
Secondary educ.	-63.29	-57.47	1.95	-7.78
	[17.26]***	[15.42]***	[0.99]**	[4.50]*
Tertiary educ.	-88.27	-87.52	1.76	-2.51
	[16.16]***	[14.38]***	[0.82]**	[4.61]
Constant	523.36	437.00	6.34	80.01
	[46.30]***	[39.70]***	[2.29]***	[13.09]***
Mean of Y	697.15	482.83	20.01	194.31
N	4671	4671	4671	4671
R^2	0.11	0.12	0.04	0.03

Notes: Dependent variable is average time spent per week on all or on a specified category of websites, measured in minutes. All equations include controls for other demographic characteristics and occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 9: Robustness checks: excluding time spent on online games and gambling

	All Websites		Leisure Websites	
	All	Excluding online games and gambling	All Leisure	Excluding online games and gambling
Income 18-27000	-50.48	-39.25	-43.04	-31.81
	[7.80]***	[6.58]***	[6.75]***	[5.38]***
Income 27-36001	-65.76	-50.83	-59.91	-44.98
	[8.21]***	[6.92]***	[7.04]***	[5.60]***
Income 36-54000	-99.47	-77.42	-88.55	-66.50
	[7.87]***	[6.65]***	[6.68]***	[5.30]***
Income 54-72000	-122.48	-94.88	-107.49	-79.89
	[9.11]***	[7.89]***	[7.55]***	[6.15]***
Income >72000	-148.03	-114.40	-123.43	-89.80
	[10.41]***	[9.17]***	[8.57]***	[7.16]***
Secondary educ.	17.71	20.73	6.34	9.37
	[7.23]**	[6.01]***	[6.18]	[4.83]*
Tertiary educ.	-1.85	10.37	-16.66	-4.45
	[6.75]	[5.60]*	[5.75]***	[4.46]
Constant	398.29	356.41	334.40	292.52
	[21.19]***	[18.48]***	[18.05]***	[15.07]***
Mean of Y	384.22	174.28	324.03	266.25
N	18680	18680	18680	18680
R^2	0.10	0.09	0.10	0.10

Notes: Dependent variable is average time spent per week on all or on a specified category of websites, measured in minutes. All equations include controls for other demographic characteristics and occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 10: Country specific results for all websites

	DE	FR	IT	ES	UK
Income 18-27000	-82.32	-42.14	-42.39	-38.56	-53.68
	[18.80]***	[18.03]**	[16.78]**	[14.83]***	[19.61]***
Income 27-36001	-79.88	-67.94	-76.66	-36.95	-77.53
	[19.48]***	[17.65]***	[17.36]***	[16.85]**	[23.50]***
Income 36-54000	-130.72	-95.86	-112.74	-73.94	-98.32
	[18.83]***	[17.13]***	[16.78]***	[16.22]***	[20.87]***
Income 54-72000	-141.96	-111.19	-88.65	-84.92	-163.38
	[21.51]***	[18.21]***	[22.43]***	[21.52]***	[23.01]***
Income >72000	-188.31	-128.99	-131.47	-130.32	-145.74
	[23.16]***	[19.36]***	[25.45]***	[24.53]***	[34.09]***
Secondary educ.	-20.76	17.52	33.44	69.09	34.43
	[16.00]	[14.25]	[16.39]**	[16.37]***	[28.92]
Tertiary educ.	-50.44	0.02	30.33	40.51	15.03
	[14.98]***	[11.19]	[18.17]*	[15.10]***	[27.03]
Constant	435.59	269.10	369.71	361.08	478.07
	[43.82]***	[31.20]***	[34.63]***	[34.22]***	[46.02]***
Mean of Y	344.77	218.86	305.93	315.19	354.14
N	3928	4028	3535	3767	3422
R^2	0.10	0.11	0.08	0.08	0.09

Notes: Dependent variable is average time spent per week on all websites, measured in minutes. All equations include controls for other demographic characteristics, occupation and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 11: Country specific results for leisure websites

	DE	FR	IT	ES	UK
Income 18-27000	-60.13	-28.69	-35.54	-39.47	-50.57
	[16.61]***	[14.95]*	[14.53]**	[12.75]***	[16.77]***
Income 27-36001	-67.79	-58.65	-66.81	-40.45	-69.37
	[17.20]***	[14.40]***	[14.84]***	[14.31]***	[20.26]***
Income 36-54000	-110.02	-77.41	-99.08	-73.70	-92.28
	[16.41]***	[13.89]***	[14.22]***	[13.62]***	[17.63]***
Income 54-72000	-128.78	-89.80	-83.66	-79.16	-138.93
	[17.82]***	[14.60]***	[18.36]***	[18.06]***	[19.35]***
Income >72000	-153.64	-103.18	-112.71	-108.62	-126.85
	[19.44]***	[15.48]***	[20.17]***	[20.82]***	[27.81]***
Secondary educ.	-24.11	9.82	11.72	51.42	17.05
	[13.68]*	[11.80]	[14.30]	[14.13]***	[24.94]
Tertiary educ.	-49.35	-11.54	5.10	19.95	-9.44
	[12.74]***	[9.13]	[15.71]	[13.03]	[23.18]
Constant	364.76	215.00	295.56	283.89	409.66
	[38.46]***	[24.92]***	[29.74]***	[28.35]***	[39.28]***
Mean of Y	240.932	141.75	222.21	230.71	241.52
N	3928	4028	3535	3767	3422
R^2	0.11	0.09	0.08	0.08	0.11

Notes: Dependent variable is average time spent per week on leisure websites, measured in minutes. All equations include controls for other demographic characteristics, occupation and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 12: Country specific results for human capital websites

	DE	FR	IT	ES	UK
Income 18-27000	-2.27	-3.31	-1.68	-0.33	-1.98
	[0.81]***	[1.35]**	[0.75]**	[0.84]	[1.21]
Income 27-36001	-2.31	-2.81	-1.53	-0.55	-2.79
	[0.75]***	[1.38]**	[0.82]*	[1.03]	[1.44]*
Income 36-54000	-2.80	-4.00	-1.72	-1.98	-1.14
	[0.88]***	[1.36]***	[0.93]*	[0.90]**	[1.38]
Income 54-72000	-3.70	-3.03	-0.15	-1.68	-3.71
	[1.00]***	[1.62]*	[1.39]	[1.66]	[1.48]**
Income >72000	-4.48	-4.36	-1.49	-4.13	-4.53
	[1.32]***	[1.54]***	[1.85]	[1.18]***	[1.56]***
Secondary educ.	0.85	2.21	2.20	3.91	3.11
	[0.57]	[1.04]**	[0.59]***	[0.84]***	[1.47]**
Tertiary educ.	2.16	2.30	3.91	4.86	3.92
	[0.80]***	[0.72]***	[0.73]***	[0.71]***	[1.27]***
Constant	8.45	4.08	3.85	5.59	1.38
	[2.26]***	[2.30]*	[1.54]**	[2.29]**	[3.12]
Mean of Y	7.41	8.91	7.71	9.90	10.12
N	3928	4028	3535	3767	3422
R^2	0.04	0.06	0.04	0.04	0.02

Notes: Dependent variable is average time spent per week on human capital websites, measured in minutes. All equations include controls for other demographic characteristics and occupation, and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 13: Country specific results for goods and services websites

	DE	FR	IT	ES	UK
Income 18-27000	-19.93	-10.14	-5.16	1.24	-1.13
	[5.19]***	[4.98]**	[4.44]	[3.93]	[5.83]
Income 27-36001	-9.78	-6.48	-8.31	4.06	-5.36
	[5.80]*	[5.04]	[4.77]*	[4.32]	[6.94]
Income 36-54000	-17.91	-14.45	-11.95	1.74	-4.90
	[5.49]***	[4.83]***	[4.74]**	[4.74]	[6.36]
Income 54-72000	-9.47	-18.37	-4.84	-4.08	-20.73
	[7.64]	[5.41]***	[7.04]	[5.59]	[6.88]***
Income >72000	-30.19	-21.44	-17.28	-17.57	-14.37
	[7.28]***	[5.63]***	[8.09]**	[5.90]***	[10.65]
Secondary educ.	2.50	5.49	19.52	13.77	14.28
	[4.90]	[4.08]	[4.09]***	[3.96]***	[7.53]*
Tertiary educ.	-3.25	9.26	21.31	15.71	20.56
	[5.04]	[3.41]***	[4.77]***	[3.80]***	[7.05]***
Constant	62.38	50.02	70.31	71.60	67.03
	[11.54]***	[10.01]***	[9.48]***	[9.67]***	[13.21]***
Mean of Y	96.43	68.19	76.02	74.57	102.49
N	3928	4028	3535	3767	3422
R^2	0.04	0.09	0.05	0.04	0.04

Notes: Dependent variable is average time spent per week on goods & services websites, measured in minutes. All equations include controls for other demographic characteristics, occupation and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets.*, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 14: Baseline model: comparison of working and not working internet users.

	All Websites		Leisure Website		Human Capital Websites		Goods & Services Websites	
	empl.	not empl.	empl.	not empl.	empl.	not empl.	empl.	not empl.
Income 18-27000	-50.70	-45.02	-45.30	-36.60	-0.95	-2.37	-4.45	-6.05
	[10.04]***	[12.87]***	[8.72]***	[11.06]***	[0.55]*	[0.64]***	[2.78]	[3.52]*
Income 27-36001	-55.20	-83.44	-53.99	-71.95	-0.65	-2.98	-0.56	-8.52
	[10.55]***	[13.71]***	[9.09]***	[11.68]***	[0.59]	[0.70]***	[2.98]	[3.84]**
Income 36-54000	-97.23	-101.41	-90.86	-84.76	-1.17	-3.29	-5.20	-13.36
	[9.95]***	[14.06]***	[8.48]***	[11.92]***	[0.54]**	[0.84]***	[2.92]*	[4.07]***
Income 54-72000	-121.77	-118.78	-110.98	-100.15	-1.30	-4.16	-9.49	-14.47
	[11.39]***	[16.69]***	[9.51]***	[13.71]***	[0.75]*	[1.03]***	[3.58]***	[5.09]***
Income >72000	-154.44	-121.86	-134.28	-94.16	-2.85	-4.27	-17.30	-23.43
	[12.29]***	[21.93]***	[10.17]***	[18.14]***	[0.67]***	[1.53]***	[3.89]***	[5.77]***
Secondary educ.	-1.94	37.17	-9.12	22.76	1.81	2.58	5.37	11.82
	[9.34]	[11.76]***	[7.96]	[10.11]**	[0.49]***	[0.59]***	[2.69]**	[3.16]***
Tertiary educ.	-22.71	18.28	-31.14	-2.26	2.21	4.62	6.22	15.92
	[8.95]**	[10.78]*	[7.63]***	[9.17]	[0.45]***	[0.58]***	[2.57]**	[3.19]***
Constant	415.33	482.58	343.29	417.67	6.01	8.84	66.02	56.06
	[25.35]***	[36.38]***	[21.25]***	[31.70]***	[1.39]***	[1.85]***	[7.06]***	[9.04]***
Mean of Y	284.73	347.18	195.69	249.53	8.25	9.83	80.79	87.82
N	12311	6369	12311	6369	12311	6369	12311	6369
R^2	0.09	0.10	0.09	0.10	0.03	0.05	0.04	0.08

Notes: Dependent variable is average time spent online on all or on a specified category of websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Heteroskedasticity robust standard errors are in brackets. *, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 15: Quantile regressions: all websites

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18-27000	-8.17	-26.95	-62.21	-86.49	-82.65	-50.48
	[2.37]***	[5.39]***	[9.40]***	[14.25]***	[26.18]***	[7.80]***
Income 27-36001	-11.63	-35.80	-70.40	-102.03	-135.86	-65.76
	[2.42]***	[5.63]***	[8.66]***	[14.01]***	[26.68]***	[8.21]***
Income 36-54000	-14.15	-44.39	-103.37	-141.80	-198.08	-99.47
	[2.37]***	[4.93]***	[8.55]***	[13.96]***	[26.52]***	[7.87]***
Income 54-72000	-15.93	-51.99	-118.13	-174.02	-260.99	-122.48
	[2.55]***	[5.62]***	[9.59]***	[15.25]***	[29.25]***	[9.11]***
Income >72000	-19.29	-63.28	-132.38	-206.48	-290.30	-148.03
	[2.66]***	[5.66]***	[11.41]***	[16.43]***	[36.96]***	[10.41]***
Secondary educ.	5.71	14.57	18.77	30.16	23.30	17.71
	[1.56]***	[3.39]***	[7.55]**	[13.02]**	[24.07]	[7.23]**
Tertiary educ.	4.83	15.59	23.62	2.86	-48.10	-1.85
	[1.47]***	[3.07]***	[7.08]***	[12.83]	[23.55]**	[6.75]
Constant	40.19	138.52	343.44	521.07	888.79	398.29
	[25.31]	[22.76]***	[68.57]***	[76.88]***	[185.45]***	[21.19]***
N	18680	18680	18680	18680	18680	18680
Pseudo R^2	0.02	0.04	0.06	0.08	0.10	0.10

Notes: Dependent variable is average time spent per week on all websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets.*, ** and *** indicate significance at 10%, 5% and 1%, respectively.

Table 16: Quantile regressions: leisure websites

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18-27000	-4.68	-17.08	-50.55	-76.44	-103.85	-4.68
	[1.41]***	[3.15]***	[6.79]***	[13.43]***	[26.08]***	[1.41]***
Income 27-36001	-6.33	-23.20	-65.46	-106.45	-144.41	-6.33
	[1.34]***	[3.05]***	[6.75]***	[12.51]***	[26.75]***	[1.34]***
Income 36-54000	-7.37	-28.02	-81.86	-145.23	-211.52	-7.37
	[1.28]***	[2.88]***	[6.21]***	[12.38]***	[25.47]***	[1.28]***
Income 54-72000	-8.63	-32.99	-92.40	-166.67	-261.85	-8.63
	[1.28]***	[3.11]***	[6.40]***	[13.62]***	[27.60]***	[1.28]***
Income >72000	-9.81	-37.26	-101.24	-175.13	-280.67	-9.81
	[1.32]***	[3.13]***	[6.83]***	[13.63]***	[31.24]***	[1.32]***
Secondary educ.	2.32	4.25	15.72	11.92	-4.95	6.34
	[0.83]***	[1.97]**	[4.81]***	[9.68]	[25.37]	[6.18]
Tertiary educ.	1.99	5.97	9.33	-13.03	-70.66	-16.66
	[0.70]***	[1.64]***	[4.11]**	[8.82]	[24.20]***	[5.75]***
Constant	22.24	95.13	235.01	494.03	663.89	334.40
	[9.64]**	[15.50]***	[52.15]***	[47.92]***	[153.08]***	[18.05]***
N	18680	18680	18680	18680	18680	18680
Pseudo R^2	0.01	0.03	0.06	0.08	0.11	0.10

Notes: Dependent variable is average time spent per week on leisure websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets. *, ** and *** indicate significance at10%, 5% and 1%, respectively.

Table 17: Quantile regressions: human capital websites

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18-27000	-0.07	-0.30	-0.77	-1.74	-4.13	-1.62
	[0.03]**	[0.07]***	[0.16]***	[0.42]***	[0.98]***	[0.42]***
Income 27-36001	-0.10	-0.36	-0.81	-2.17	-4.04	-1.64
	[0.03]***	[0.07]***	[0.18]***	[0.46]***	[1.31]***	[0.46]***
Income 36-54000	-0.16	-0.47	-1.05	-2.59	-5.15	-2.14
	[0.03]***	[0.08]***	[0.17]***	[0.43]***	[1.08]***	[0.46]***
Income 54-72000	-0.16	-0.55	-1.49	-3.85	-8.03	-2.52
	[0.03]***	[0.08]***	[0.19]***	[0.51]***	[1.26]***	[0.60]***
Income >72000	-0.22	-0.72	-2.04	-4.70	-9.22	-3.80
	[0.03]***	[0.08]***	[0.19]***	[0.57]***	[1.34]***	[0.63]***
Secondary educ.	0.11	0.30	0.89	2.15	4.06	2.33
	[0.02]***	[0.05]***	[0.11]***	[0.28]***	[0.78]***	[0.37]***
Tertiary educ.	0.16	0.45	1.44	3.88	7.44	3.39
	[0.02]***	[0.05]***	[0.12]***	[0.31]***	[0.72]***	[0.35]***
Constant	0.17	0.95	1.92	8.54	22.71	5.77
	[0.37]	[0.40]**	[0.90]**	[4.97]*	[30.78]	[1.15]***
N	18680	18680	18680	18680	18680	18680
Pseudo R^2	0.01	0.01	0.03	0.04	0.05	0.03

Notes: Dependent variable is average time spent per week on human capital websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets. *, ** and *** indicate significance at10%, 5% and 1%, respectively.

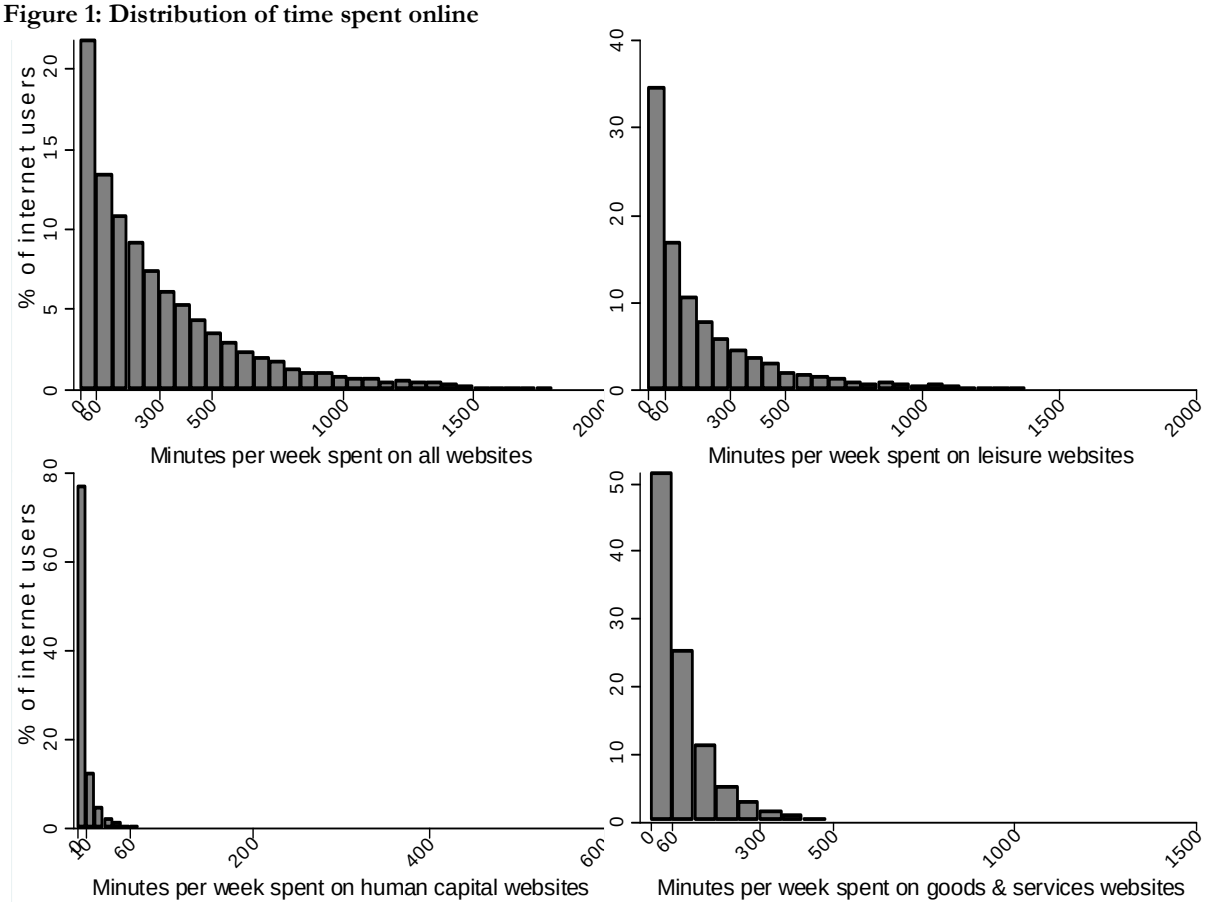
Table 18: Quantile regressions: goods & services websites

	Q10	Q25	Q50	Q75	Q90	Mean(OLS)
Income 18-27000	-0.27	-1.73	-3.36	-5.95	-12.96	-5.82
	[0.58]	[1.23]	[1.74]*	[3.16]*	[7.04]*	[2.16]***
Income 27-36001	-0.88	-1.92	-2.81	-1.86	-4.58	-4.20
	[0.60]	[1.25]	[1.94]	[3.51]	[7.14]	[2.31]*
Income 36-54000	-1.59	-4.59	-7.84	-11.19	-13.85	-8.78
	[0.56]***	[1.14]***	[1.94]***	[3.68]***	[6.86]**	[2.31]***
Income 54-72000	-1.69	-5.35	-10.06	-20.36	-27.13	-12.47
	[0.65]***	[1.37]***	[2.26]***	[3.93]***	[9.18]***	[2.85]***
Income >72000	-2.90	-9.73	-16.14	-25.07	-41.38	-20.80
	[0.62]***	[1.56]***	[2.72]***	[4.65]***	[8.97]***	[3.14]***
Secondary educ.	2.04	4.79	10.23	14.20	18.74	9.03
	[0.43]***	[0.90]***	[1.53]***	[3.31]***	[5.87]***	[2.01]***
Tertiary educ.	2.37	6.46	13.36	17.46	26.35	11.43
	[0.44]***	[0.96]***	[1.45]***	[3.05]***	[5.34]***	[1.94]***
Constant	5.76	29.10	35.78	74.95	147.79	54.62
	[4.34]	[10.55]***	[19.51]*	[23.23]***	[39.98]***	[5.58]***
N	18680	18680	18680	18680	18680	18680
Pseudo R^2	0.03	0.04	0.04	0.04	0.05	0.05

Notes: Dependent variable is average time spent per week on goods & services websites, measured in minutes. All equations include controls for other demographic characteristics, occupation, country and country-region fixed effects. Bootstrapped standard errors are in brackets. *, ** and *** indicate significance at10%, 5% and 1%, respectively.

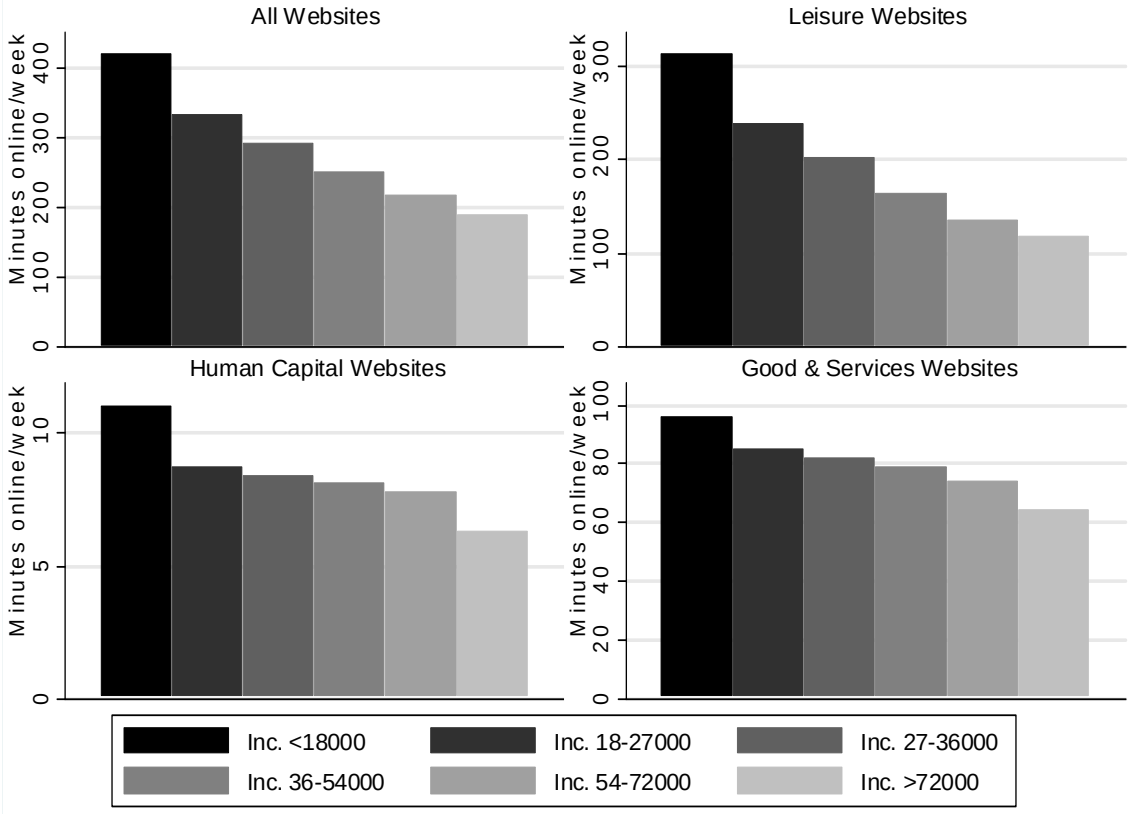
Figures

Figure 1: Distribution of time spent online



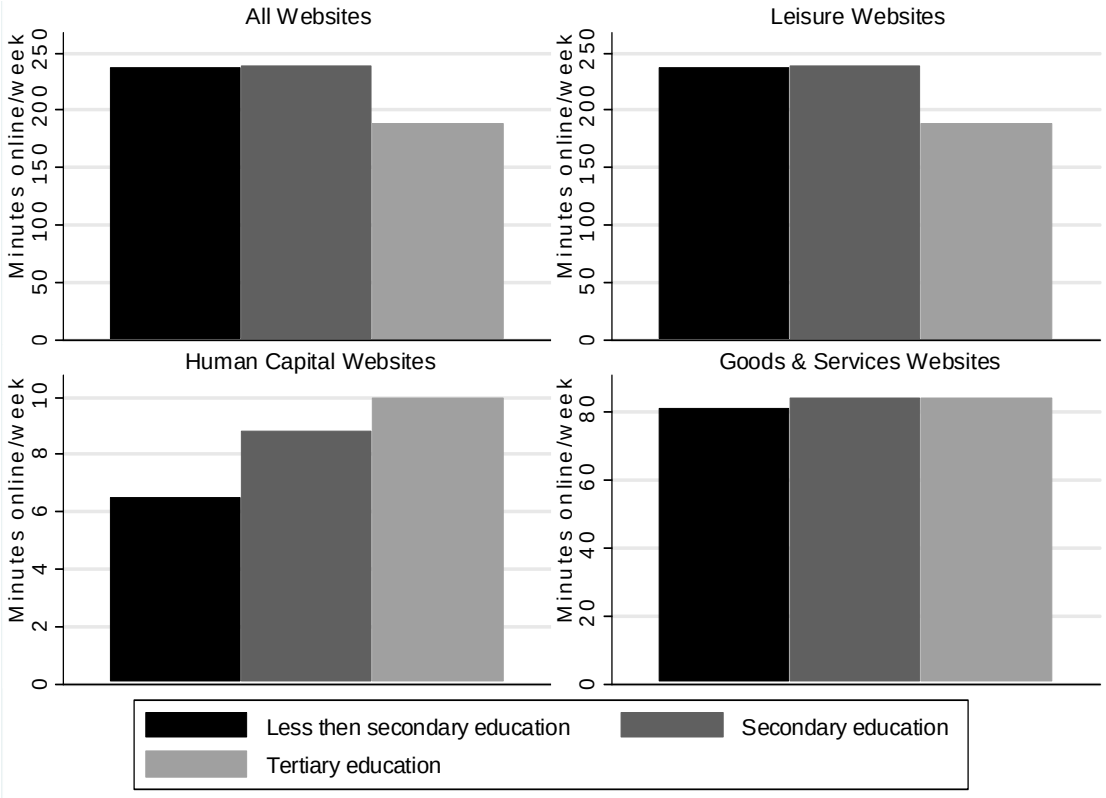
Source: Calculations based on Nielsen Click stream

Figure 2: Time spent on different websites and household income



Source: Calculations based on Nielsen Click stream

Figure 3 Time spent on different websites and education



Source: Calculations based on Nielsen Click stream

European Commission
EUR 25959 – Joint Research Centre – Institute for Prospective Technological Studies

Title: Has the Digital Divide Been Reversed? Evidence from Five EU Countries

Authors: Smaranda Pantea, Bertin Martens

Luxembourg: Publications Office of the European Union

2013 – 30 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)

ISBN 978-92-79-29822-6 (pdf)

doi:10.2791/12705

Abstract
This paper examines the digital divide in internet use in general and checks whether there may be digital divides in internet use for specific purposes (leisure, improving human capital and obtaining goods and services). It uses a unique dataset which covers the entire clickstream of almost 20.000 internet users in the five largest EU economies during 2011. Our main finding is that, for those who have access to the Internet, the digital divide in internet use has been reversed. Low-income internet users spend more time on the internet than high-income users. In addition, we find the effect of income on internet use is not affected by employment status of the internet users and we discuss several possible explanations for this result. There is some evidence of an education-based digital divide in the ability to use websites related to career, education, health and buying and obtaining goods and services. Tertiary education has a negative effect on time spent on leisure websites and a positive effect on time spent on websites related to human capital and obtaining services and goods. Using quantile regressions, we find that the negative effect of income on time spent on internet and the positive effect of education on time spent on websites related to career, education, health and buying and obtaining goods and services hold for the entire conditional distribution of these online activities and that these effects are higher for the upper quantiles than for the lower quantiles.
JEL codes: L86, D12, D13
Keywords: Internet Use, Time allocation, Leisure.

LF-NA-25959-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

ISBN 978-92-79-29822-6

9 789279 298226